 DOI 10.1515/iwp-2012-0059 — Information. Wissenschaft & Praxis 2012; 63(5):314–318

David Elsweiler, Regensburg

Understanding the Complexities of Email Behaviour

This position statement summarises three studies from a project aiming to learn about and support email search behaviour. The findings combine to form a rich and multifaceted picture of user behaviour and demonstrate why it is important to account for user behaviour at all stages of an Information Science project.

Keywords: information behaviour, empirical study

Informationsverhalten und E-Mails. Komplexe Verhaltensweisen verstehen
Dieser kurze Forschungsbericht stellt drei Studien vor, die sich mit dem Informationsverhalten von Personen im Kontext von E-Mails befassen. Die Ergebnisse dieser Studien zeigen, dass diese Verhaltensweisen komplex und facettenreich sind. Umso mehr lohnt es, im Rahmen informationswissenschaftlicher Forschung das Verhalten von Benutzerinnen und Benutzern in den Mittelpunkt zu stellen.

Deskriptoren: Informationsverhalten, empirische Untersuchung

L'usage de l'information et le courriel. Comprendre une manière d'agir multiple
L'article analyse trois études concernant l'usage du courriel. Les résultats montrent, que les manières d'utilisation sont complexes et multiples. Par conséquent l'auteur souligne que l'étude des besoins des utilisateurs est un pilier central pour chaque projet de recherche dans le domaine des sciences de l'information.

Mots-clés: enquête, courrier électronique, utilisation, recherche empirique

1 Introduction

Information behaviour refers to the many ways in which humans interact with information and information objects. This includes, but is not restricted to, the processes of creating, seeking, acquiring, organising, sharing and using information (Wilson, 2000). Information behaviour is complex and depends not only on tools available to the user, but the user's socio-cognitive skills, as well as a complicated array of contextual and task factors (Ingwersen & Järvelin, 2005). Research projects at the Chair for Information Science at the University of Regensburg are helping to shed light on how people behave and why, as well as how information systems should be designed to reflect the needs and behaviour of the user.

Three projects with which I am personally involved are contributing to our knowledge of how people behave with information. These projects are investigating behaviour with email messages, information behaviour in non-work situations (Elsweiler, Wilson & Kirkegaard Lunn, 2011) and the personalised recommendation of recipes (see page 325–329). While these are a diverse set of topics, they are related in that they combine to show the variety and complexity of information behaviour and provide a basis from which we can study factors that influence how a user of an information system behaves.

In this short position statement, for space reasons, the focus will be on email behaviour. Of the three projects, our email research is the most mature. By summarising some of the work already performed it will allow us to showcase our approach to understanding information behaviour, which is to use several different types of investigation and combine the findings to form a richer picture of behaviour than would be possible using any single method in isolation. The final section will conclude the article by relating the described work to the other projects and outline my thoughts on what the findings so far mean for studying information science.

2 Email Search Behaviour

Despite the recent explosive growth of social media applications, email remains the most popular communication medium in use today. An estimated 294 billion emails are sent each day (Radicati, 2010), dwarfing the 1 billion facebook entries[1] and 200 million tweets[2] which are posted monthly. Email is, however, much more than just a communication tool. People use email for diverse purposes including the management of tasks, projects, contacts and documents (Whittaker & Sidner, 1996). Email is also not the ephemeral media it was originally intended to be. Most messages have a lifespan of several weeks or months and some messages are re-read years after they were first received (Elsweiler, Losada, Toucedo & Fernandez, 2011).

Reflecting the diverse usage patterns and long-life spans, studies of desktop search logs show that email messages tend to be searched for more often than any other kind of media, including visited web pages (Dumais et al., 2003), and re-finding information that has been seen or used before is known to be a challenging task for people generally (Boardman & Sasse, 2004). We have been using a variety of methods in an attempt to achieve a holistic understanding of how people look for and use information in their email collections and how email systems could be designed to support these activities.

3 A Laboratory Study

A lab-based study allowed the investigation the recollections people have for the information they are trying to find (Elsweiler, Baillie & Ruthven, 2008). Recollections are important when searching for emails because the user knows that the mail exists – he has seen it before -and the recollection of the mail's content and the context surrounding it being sent is what guides the information seeking process (Capra & Perez-Quinones, 2005).

Our results show that people tend to have good recollection for their email messages, remembering a wide-range of contextual factors associated with messages (when the email was sent, who sent it, why it was sent, etc.). However, there is a steep drop-off in the extent and quality of recollection after very short time periods. Some users are more susceptible to loss of recollection than others. For example, in our study, the users who received the most messages demonstrated a dramatic reduction in recollection abilities compared to other participants after only 7 days. The filing strategy employed by a user is also highly predictive of the quality of recollections for an email. Our data show that people who file their emails ("filers") tended to have much poorer recollection than those who make little or no effort to organise their mails. While it may initially seem strange that those who exert cognitive effort to decide where a message should be stored should remember the less than those who do not, it seems that moving emails into a folder brings about "out of sight, out of mind" problem (Bruce, Jones & Dumais, 2004). When mails are in a folder, the user is less likely to interact with the message and there is no reinforcement function. Other analyses provide further evidence that email systems should better support people who choose to file. Filers tend to perceive tasks as more difficult (Elsweiler, Baillie & Ruthven, 2011) and are less successful at finding what they are looking for (Elsweiler, Baillie & Ruthven, 2009).

4 A Naturalistic Study

Results from a second study provide insight into the reasons behind the findings reported in the previous section. We collected data via a naturalistic study of email use with Mozilla Thunderbird. The aim was to complement the laboratory study by learning how and how often people re-use email messages in the wild.

We developed and deployed a custom software extension that recorded user interactions with the client including messages that were read, clicks on folders, clicking on column headers to sort mails and search queries submitted (Elsweiler, Harvey & Hacker, 2011). A large ($n = 47$) and diverse population provided 4 months of interaction data. After cleaning, these data provide rich information about how and how often people look for emails in their collections. The vast majority (93.8%) of tasks involved a query and queries tend to be very short. Most (90.88%) were single words and only 0.8% had a length greater than 2 terms. The queries submitted very often contained references to people with most searches heavily featuring the people who send the most emails to the user. The top 10 senders and the top 5 in particular were searched for far more often than would be expected based on the frequency with which they send mails alone (Harvey & Elsweiler, 2012).

1 http://www.facebook.com/press/info.php?statistics [13.9.2011].
2 http://blog.twitter.com/2011/06/200-million-tweets-perday.html [13.9 2011].

Regarding the frequency re-finding our participants can be clustered into two groups: one group ($n = 14$) performed on average over 2 tasks per day and a second group ($n = 33$) performed 1 task roughly every 4 days. The long-tailed distribution pervaded many of the analyses performed. For example, most tasks were short (median = 4 mins), contained few message views (median = 4) and involved recently viewed messages, but there are also tasks which took much longer (max = 72), contained many more message views (max = 290) and involved older mails (max = 6.5 yrs). The data mirror many of the findings from the lab-based study. For example, users with large numbers of folders often took longer and required more interactions to locate messages. In such tasks, the participant regularly selected multiple folders before the task was completed, suggesting that users are not always good at remembering the correct folder. Further, in long re- finding tasks, filers often look in the same folder multiple times, suggesting disorientation. These are two possible explanations for the repeated finding that filers have difficulties when re-finding.

5 Simulated experiments

A third aspect of our research relating to email behaviour is our work on the simulated evaluation of personal search. Personal search, which includes email search, is a system that allows the user to locate or re-find information items they have seen, created or used at some point in the past (Elsweiler, Jones, Kelly & Teevan, 2010). Evaluating retrieval algorithms or interfaces for personal search is notoriously difficult and few evaluations have been performed (Kelly & Teevan, 2007). To test retrieval models in Cranfield-style evaluations, researchers use public collections that have some of the properties of a personal collection. For email, this tends to be messages sent and received by active members of the W3C mailing list (Kim & Croft, 2009). Search tasks are simulated by randomly selecting items from the collection to create known-item retrieval tasks with systems tested on their ability to locate these documents from the collections with simulated queries. Simulated queries are generated by drawing terms from the target document to fit a query length distribution.

This is a potentially powerful method of scientific evaluation for Personal Search, which, in contrast to user studies, is low cost and easily repeatable. There are, however, problems with current implementations. These are over-simplified and make assumptions about user behaviour that are unlikely to be true. For example, query terms are typically drawn independently from the document and either do not make use of field information (Azzopardi, Rijke & Balog, 2007) or assume that all fields are equally likely to be queried on (Kim & Croft, 2009).

Furthermore, current implementations do not incorporate what we already know about user behaviour e.g. that people often make use of named entities in queries. The findings of the laboratory and log-based studies described above suggest that the kinds of queries submitted will change in different scenarios. As re-finding behaviour is guided by user recollections and people remember different things in different situations, with this being heavily influenced by contextual factors, it was our hypothesis that the properties of queries will change in different situations.

To test this hypothesis we analysed the queries logged in the laboratory study described above to understand if and how the properties of queries change in various situations. We developed a number of logistic regression models to understand the influence various contextual factors had on the characteristics of submitted queries. Logistic regression is a useful way of describing the relationship between one or more independent variables (e.g., the number of emails in a collection or the user filing strategy) and a binary response variable that has only two possible values expressed as a probability, such as ("contains a Named Entity (NE)" or "does not contain a NE"). We were interested in several query characteristics including, the field the query was applied to (e.g. "contains a clause for Sender field"), types of named entities contained within the query and query length (i.e, whether the query is longer or shorter than the mean value).

The generated models indicate that several variables had an influence on the users' querying behaviour. For example we found that:

- Participants with older collections tended to submit longer queries and were more likely to query on the subject field. The age of the collection also seemed to influence the way NEs were used in the queries. For example, queries submitted against older collections were much more likely to contain a NE than those submitted against newer collections.
- Experienced participants were more likely to query on the subject field than less-experienced participants, but were less likely to query on the sender field.
- When the participants were looking for older information they were much more likely to query on the body field. When they were looking for newer information the queries submitted were less likely to con-

tain a reference to a person, but more likely to include other NEs such as organisations, books, groups or programming languages. This situation was reversed for older information with queries submitted in these situations more likely to contain references to people.
- There was a correlation between the task difficulty rating and the probability of the query including a reference to a person with easier tasks more likely to have such a reference. Different tasks were much less likely to be on the sender field.
- The filing strategy that the participant applied also seemed to influence their querying behaviour. Participants with more than 1 folder were much less likely to query on the body of an email than those who kept all of their messages in the inbox. Further, of those participants who did use folders, those who tended to tidy up their folders occasionally (spring-cleaners), were less likely to query on the body of the email than those who filed emails regularly.

The main conclusions from these analyses are that: 1) querying behaviour changes in different situations, and 2) clear relationships exist between certain variables and the query submitted in terms of length, field to which the terms were submitted and the use of NEs in the query.

Extending this work, we used the statistics for queries submitted in different situations to seed the query simulation process. We generated a set of queries for various ages of information, levels of user expertise, filing strategy and collection age. Using these queries, we tested a number of standard retrieval models to see how performance varied for different profiles. The experiments performed showed that different retrieval models provided best performance in different scenarios. This not only demonstrates the utility of seeding the evaluation process with real data, but perhaps also suggests that search applications should directly adapt their behaviour depending on the context.

6 Summary and Conclusions

In this article we have briefly summarised three studies related to email re-finding behaviour, highlighting some of the key findings. This work shows how an enhanced understanding of a topic can be achieved by combining the findings of multiple kinds of study.

The findings of these studies hint at several design implications, for example, highlighting that particular tasks and users need to be better supported. Filers have difficulties because they tend to remember less about their messages, perhaps because they are less likely to interact with messages after they have been put in a folder. They also have difficulties later identifying the correct folder to look in and can become disorientated in larger folder structures. Specific features might be designed to assist with these problems. Similarly, we know that users tend to repeatedly re-find messages from the same senders. An interface that filters messages for popular senders may then save users lots of time. These are only two of many design implications highlighted based on the findings. Many more are listed and described in detail in the cited publications.

My belief is that the findings also have implications for the way we should approach the study of Information Science. What these email studies show is that information systems will not always be used as they are designed. Users find inventive uses to suit their own needs and their behaviour with the system will reflect this. Our studies also demonstrate that the way the system is used will influence not only the user's success, but also other cognitive aspects, such as how they perceive or remember information. Contextual variables are vitally important as we have seen that such cognitive aspects and resultant behaviour (queries) can change drastically in different situations and accounting for this in systems evaluations heavily influences the outcomes.

Other information behaviour projects at the Chair for Information Science support these arguments further. For example, our work on casual-leisure search (Elsweiler, Wilson, & Kirkegaard Lunn, 2011) has revealed strong differences in user aims and behaviour in search tasks in non-work situations compared to the work tasks normally studied. Our healthy recipe recommendation work goes even further; the premise being that user behaviour can be positively changed by providing appropriate information (see page 325–329). All of this work underlines the importance of accounting for user behaviour at all stages of the design process and I believe, endorses a behavioral perspective to studying information science (Ellis, 1989).

References

Azzopardi, L.; Rijke, M. de & Balog, K. (2007). Building simulated queries for known-item topics: an analysis using six european languages. In: Sigir '07: Proceedings of the 30th annual international acm sigir conference on research and development in information retrieval (pp. 455–462). New York, NY, USA: ACM.

Boardman, R. & Sasse, M. A. (2004). “stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management. In: E. Dykstra-Erickson & M. Tscheligi (Eds.): Chi ’04: Proceedings of the sigchi conference on human factors in computing systems (pp. 583–590). New York, NY, USA: ACM Press.

Bruce, H.; Jones, W. & Dumais, S. (2004, October). Keeping and refinding information on the web: What do people do and what do they need? In: Asist 2004.

Capra, R. G. & Perez-Quinones, M. A. (2005). Using web search engines to find and refind information. In: Computer 38(10), 36–42.

Dumais, S.; Cutrell, E.; Cadiz, J.; Jancke, G.; Sarin, R. & Robbins, D. (2003). Stuff i’ve seen: a system for personal information retrieval and re-use. In: Proc. acm sigir ’03 (pp. 72–79).

Ellis, D. (1989). A behavioural model for information retrieval system design. In: Journal of Information Science 15(4), 237–247.

Elsweiler, D.; Baillie, M. & Ruthven, I. (2008). Exploring memory in email refinding. In: ACM Trans. Inf. Syst. 26(4), 1–36.

Elsweiler, D.; Baillie, M. & Ruthven, I. (2009). On understanding the relationship between recollection and refinding. In: Journal of Digital Information (JoDI) 10(5).

Elsweiler, D.; Baillie, M. & Ruthven, I. (2011). What makes re-finding information difficult? a study of email re-finding. In: Proc. ecir 2011.

Elsweiler, D.; Harvey, M. & Hacker, M. (2011). Understanding refinding behavior in naturalistic email interaction logs. In: Proc. acm sigir (pp. 35–44). ACM. Available from http://doi.acm.org/10.1145/2009916.2009925

Elsweiler, D.; Harvey, M. & Ludwig, B. (2012). Kochen im Kontext – Am Alltag orientierte Empfehlung von Gerichten. In: Information. Wissenschaft und Praxis 63(5), XX–XX.

Elsweiler, D.; Jones, G. J. F.; Kelly, L. & Teevan, J. (Eds.). (2010). Sigir workshop on desktop search workshop: Understanding, supporting, and evaluating personal data search.

Elsweiler, D.; Losada, D. E.; Toucedo, J. & Fernandez, R. (2011). Seeding simulated queries with userstudy data forpersonal search evaluation. In: Proc. acm sigir (pp. 25–34). ACM. http://doi.acm.org/10.1145/2009916.2009924

Elsweiler, D.; Wilson, M. L. & Kirkegaard Lunn, B. (2011). 5 New directions in information behaviour. In: A. Spink & J. Heinstrm (Eds.), (chap. Understanding Casual-leisure Information Behaviour). Emerald Publishing.

Harvey, M. & Elsweiler, D. (2012). Exploring query patterns in email search. In: R. Baeza-Yates et al. (Eds.): Advances in information retrieval (Vol. 7224, p. 25-36). Berlin; Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-642-28997-2 (10.1007/978-3-642-28997-2)

Ingwersen, P. & Järvelin, K. (2005). The turn: Integration of information seeking and retrieval in context (Vol. 18). Springer. (ISBN: 1-4020-3850-X)

Kelly, D. & Teevan, J. (2007). Personal information management. In: W. Jones & J. Teevan (Eds.): Personal Information Management Understanding what works: Evaluating personal information management tools (p. 190–204). Seattle: University of Washington Press.

Kim, J. & Croft, W. B. (2009). Retrieval experiments using pseudo-desktop collections. In: Cikm ’09: Proceeding of the 18th acm conference on information and knowledge management (pp. 1297–1306). New York, NY, USA: ACM.

Radicati, S. (2010). Email statistics report. http://www.radicati.com/wp/wp-content/uploads/2010/04/Email-Statistics-Report-2010-2014-Executive-Summary2.pdf

Whittaker, S. & Sidner, C. (1996). Email overload: exploring personal information management of email. In: Sigchi 1996 (pp. 276–283).

Wilson, T. D. (2000). Human information behavior. In: Informing Science 3(2), 49–55.

Dr. David Elsweiler
Lehrstuhl für Informationswissenschaft
Universität Regensburg
Universitätsstraße 31
93053 Regensburg
david.elsweiler@ur.de

David Elsweiler has been employed at the Chair for Information Science since August 2011. Previously he was an Alexander von Humboldt Research Fellow at the University of Erlangen-Nuremberg. After completing his PhD under the supervision of Prof. Ian Ruthven in 2007 he has established himself as a prominent scholar in the information retrieval community. He regularly publishes and reviews for top publications in this area and in 2011 received an ACM SIGIR “Outstanding Paper Award”.